Notes for Asymptotic Analysis Examples

T. V. K. Chaitanya

March 23, 2011

This notes gives a quick introduction to an array processing application and a source localization application to be discussed in the lecture of asymptotic analysis of estimators.

I. Notation

- $(.)^*$: Hermitian transpose
- tr $\{\boldsymbol{A}\}$: trace of Matrix $\boldsymbol{A}$
- $\mathcal{R}\{\boldsymbol{A}\}$: range space of $\boldsymbol{A}$
- $\|\boldsymbol{A}\|_F$: Frobenius norm, $\|\boldsymbol{A}\|_F^2 = \text{tr}\{\boldsymbol{A}^*\boldsymbol{A}\}$
- $\boldsymbol{A}^\dagger$: pseudoinverse for $\boldsymbol{A}$ of full rank, $\boldsymbol{A}^\dagger = (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*$
- $\boldsymbol{P}_{\boldsymbol{A}}$: projection matrix, $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{A}^\dagger = \boldsymbol{A}(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*$
- $\boldsymbol{P}_{\boldsymbol{A}}^\perp$: $\boldsymbol{P}_{\boldsymbol{A}}^\perp = \boldsymbol{I} - \boldsymbol{P}_{\boldsymbol{A}}$

II. Array Processing Application

System model and the preliminary results for subspace fitting based array processing is presented in this section. The material presented below is taken mainly from [1].

- Sensor array with m elements
- d emitters at different locations, $d < m$.
- The output of ith sensor is

$$x_i(t) = \sum_{j=1}^{d} a_i(\theta_j)\, s_j(t) + n_i(t) \tag{1}$$

where $a_i(\theta_j)$ is a complex scalar, $s_j(t)$ is jth emitters signal and $n_i(t)$ is the noise at ith sensor.

- In vector form, we can write

$$\boldsymbol{x}(t) = \begin{bmatrix} \boldsymbol{a}(\theta_1) & \cdots & \boldsymbol{a}(\theta_d) \end{bmatrix} \begin{bmatrix} s_1(t) \\ \vdots \\ s_d(t) \end{bmatrix} + \begin{bmatrix} n_1(t) \\ \vdots \\ n_m(t) \end{bmatrix}$$
$$= \boldsymbol{A}(\boldsymbol{\theta}_0)\,\boldsymbol{s}(t) + \boldsymbol{n}(t) \quad (2)$$

  where $\boldsymbol{a}(\theta_j) = [a_1(\theta_j), \ldots, a_m(\theta_j)]^T$ is sensor response to a unit wavefront from direction θ_j. $\boldsymbol{\theta}_0$ is a d-dimensional parameter vector corresponding to true DOA's.
- $G = \{\boldsymbol{a}(\theta_j) | \theta_j \in \Theta\}$ is the array manifold. It is assumed that the parametrization of G is known.
- $G^d = \left\{ \boldsymbol{A}(\boldsymbol{\theta}) | \boldsymbol{A}(\boldsymbol{\theta}) = \begin{bmatrix} \boldsymbol{a}(\theta_1) & \cdots & \boldsymbol{a}(\theta_d) \end{bmatrix}, \theta_1 < \theta_2 < ... < \theta_d \right\}$. G^d is parametrized by $\boldsymbol{\theta} = [\theta_1 \theta_2 ... \theta_d]^T$.
- $\boldsymbol{A}(\boldsymbol{\theta}) \in G^d$ has full rank (d).
- Sampled array output can be written as

$$\boldsymbol{X}_N = [\boldsymbol{x}(1) \ldots \boldsymbol{x}(N)] = \boldsymbol{A}(\boldsymbol{\theta}_0)\,\boldsymbol{S}_N + \boldsymbol{N}_N \quad (3)$$

  where $\boldsymbol{X}_N$ is $m \times N$, $\boldsymbol{A}(\boldsymbol{\theta}_0)$ is $m \times d$, $\boldsymbol{S}_N$ is $d \times N$ and $\boldsymbol{N}_N$ is $m \times N$.
- $\mathbb{E}[\boldsymbol{n}(t)\,\boldsymbol{n}^*(t)] = \sigma^2 \boldsymbol{I}_m$.
- $\boldsymbol{S} = \mathbb{E}[\boldsymbol{s}(t)\,\boldsymbol{s}^*(t)] = \lim_{n\to\infty} \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{s}(t)\,\boldsymbol{s}^*(t)$
- $\boldsymbol{R} = \mathbb{E}[\boldsymbol{x}(t)\,\boldsymbol{x}^*(t)] = \lim_{n\to\infty} \frac{1}{N} \sum_{t=1}^{N} \boldsymbol{x}(t)\,\boldsymbol{x}^*(t) = \boldsymbol{A}(\boldsymbol{\theta}_0)\,\boldsymbol{S}\boldsymbol{A}^*(\boldsymbol{\theta}_0) + \sigma^2 \boldsymbol{I}_m$.
- $\boldsymbol{S}$ is of rank $d^{'} \leq d$. Which means that the signal waveforms are linearly dependent.
- For asymptotic analysis, the array manifold vectors are assumed to be twice continuously differentiable with bounded derivatives in a neighborhood of true DOA's.
- To enable unique identification of the signal parameters, the following condition is imposed on the array response

$$\boldsymbol{A}(\boldsymbol{\theta})\,\boldsymbol{T} = \boldsymbol{A}(\boldsymbol{\eta})\,\boldsymbol{U} \Longrightarrow \boldsymbol{\theta} = \boldsymbol{\eta}$$

  For $\boldsymbol{A} \in G^d$, this condition is met if $d < \left(m + d^{'}\right)/2$.
- Eigendecomposition of

$$\boldsymbol{R} = \sum_{i=1}^{m} \lambda_i \boldsymbol{e}_i \boldsymbol{e}_i^* = \boldsymbol{E}_s \boldsymbol{\Lambda}_s \boldsymbol{E}_s^* + \sigma^2 \boldsymbol{E}_n \boldsymbol{E}_n^* \quad (4)$$

  where $\lambda_1 > \lambda_2 > \cdots > \lambda_{d'} > \lambda_{d'+1} = \cdots = \lambda_m = \sigma^2$. $\boldsymbol{E}_s$ is $m \times d^{'}$, $\boldsymbol{E}_n$ is $m \times m - d^{'}$and $\boldsymbol{\Lambda}_s$ is

$d^{'} \times d^{'}$.

- $\mathcal{R}(\boldsymbol{E}_s) \subseteq \mathcal{R}(\boldsymbol{A}(\boldsymbol{\theta}_0))$ with equality iff $d^{'} = d$.
- The eigendecomposition of sample covariance matrix $\hat{\boldsymbol{R}} = \frac{1}{N}\boldsymbol{X}_N\boldsymbol{X}_N^* = \hat{\boldsymbol{E}}_s\hat{\boldsymbol{\Lambda}}_s\hat{\boldsymbol{E}}_s^* + \hat{\boldsymbol{E}}_n\hat{\boldsymbol{\Lambda}}_n\hat{\boldsymbol{E}}_n^*$
- It is assumed that d and $d^{'}$ are known.

A. Basic Subspace fitting problem

- We can write the basic subspace fitting problem as

$$\hat{\boldsymbol{A}}, \hat{\boldsymbol{T}} = \arg\min_{\boldsymbol{A},\boldsymbol{T}} \|\boldsymbol{M} - \boldsymbol{A}\boldsymbol{T}\|_F^2 \tag{5}$$

where $\boldsymbol{M}$ is $m \times q$ matrix and $\boldsymbol{T}$ is a $p \times q$ matrix. For a fixed $\boldsymbol{A}$, the minimum w.r.t $\boldsymbol{T}$ is a measure of how well the range spaces of $\boldsymbol{A}$ and $\boldsymbol{M}$ match. The estimate of $\boldsymbol{\theta}$ is obtained from the parameters of $\hat{\boldsymbol{A}}$.

- The problem in (5) is separable in $\boldsymbol{A}$ and $\boldsymbol{T}$ [2]. Using $\hat{\boldsymbol{T}} = \boldsymbol{A}^\dagger\boldsymbol{M}$, we can write

$$\begin{aligned}
\hat{\boldsymbol{A}} &= \arg\min_{\boldsymbol{A}} \left\|\boldsymbol{M} - \boldsymbol{A}\boldsymbol{A}^\dagger\boldsymbol{M}\right\|_F^2 \\
&= \arg\min_{\boldsymbol{A}} \|\boldsymbol{M} - \boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\|_F^2 \\
&= \arg\min_{\boldsymbol{A}} \left\|\boldsymbol{P}_{\boldsymbol{A}}^{\perp}\boldsymbol{M}\right\|_F^2 \\
&= \arg\max_{\boldsymbol{A}} \|\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\|_F^2 \\
&= \arg\max_{\boldsymbol{A}} \ \mathrm{tr}\{\boldsymbol{M}^*\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\} \\
&\overset{(a)}{=} \arg\max_{\boldsymbol{A}} \ \mathrm{tr}\{\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\boldsymbol{M}^*\}
\end{aligned} \tag{6}$$

where $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\boldsymbol{A}^\dagger$ is a projection matrix that projects onto the column space of $\boldsymbol{A}$. In (a), we have used the fact that $\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{P}_{\boldsymbol{A}}$ and that $\mathrm{tr}\{\boldsymbol{M}^*\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\} = \mathrm{tr}\{\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{M}\boldsymbol{M}^*\}$.

1) Deterministic ML:

- Treat $\{\boldsymbol{s}(t)\}_{t=1}^N$ as parameters.
- $p(\boldsymbol{X}_N|\boldsymbol{S}_N, \boldsymbol{\theta}) \propto \exp\left(-\frac{1}{\sigma^2}\|\boldsymbol{X}_N - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S}_N\|_F^2\right)$ and we have

$$\max_{\boldsymbol{S}_N,\boldsymbol{\theta}} p(\boldsymbol{X}_N|\boldsymbol{S}_N, \boldsymbol{\theta}) = \min_{\boldsymbol{S}_N,\boldsymbol{\theta}} \|\boldsymbol{X}_N - \boldsymbol{A}(\boldsymbol{\theta})\boldsymbol{S}_N\|_F^2$$

- With $\boldsymbol{M} = \frac{1}{\sqrt{N}}\boldsymbol{X}_N$ and $\boldsymbol{A} \in G^d$ in (5) gives deterministic ML estimate. With $\hat{\boldsymbol{S}}_N = \boldsymbol{A}^\dagger(\boldsymbol{\theta})\boldsymbol{X}_N$, we have equivalent problem as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} \ \text{tr}\left\{\boldsymbol{P}_{\boldsymbol{A}}\left(\boldsymbol{\theta}\right)\hat{\boldsymbol{R}}\right\} \tag{7}$$

- (7) is similar to (6) with $\hat{\boldsymbol{R}} = \boldsymbol{M}\boldsymbol{M}^*$.

2) MD-MUSIC:

- When $d^{'} < d$, one dimensional MUSIC algorithm is not optimal and an equivalent multi-dimensional MUSIC algorithm was proposed in [3].
- One can write the estimator using subspace fitting formulation as

$$\begin{aligned} \hat{\boldsymbol{\theta}} &= \arg\min_{\boldsymbol{A}\in G^d, \boldsymbol{T}} \left\|\hat{\boldsymbol{E}}_s - \boldsymbol{A}\boldsymbol{T}\right\|_F^2 \\ &= \arg\max_{\boldsymbol{A}\in G^d} \ \text{tr}\left\{\boldsymbol{P}_{\boldsymbol{A}}\left(\boldsymbol{\theta}\right)\hat{\boldsymbol{E}}_s\hat{\boldsymbol{E}}_s^*\right\} \end{aligned} \tag{8}$$

- (8) is similar to (6) with $\boldsymbol{M} = \hat{\boldsymbol{E}}_s$.

B. Asymptotic Results

Lemma 2 in [1] proves that the deterministic ML estimator in (7), can equivalently (in the asymptotic distribution sense) be written as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{A}\in G^d} \ \text{tr}\left\{\boldsymbol{P}_{\boldsymbol{A}}\left(\boldsymbol{\theta}\right)\hat{\boldsymbol{E}}_s\left(\boldsymbol{\Lambda}_s - \sigma^2\boldsymbol{I}\right)\hat{\boldsymbol{E}}_s^*\right\} \tag{9}$$

Hence for both deterministic ML and MD-MUSIC cases, we can write the parameter estimate as

$$\hat{\boldsymbol{\theta}} = \arg\max_{\boldsymbol{\theta}} V\left(\boldsymbol{\theta}\right) = \arg\max_{\boldsymbol{\theta}} \ \text{tr}\left\{\boldsymbol{P}_{\boldsymbol{A}}\left(\boldsymbol{\theta}\right)\hat{\boldsymbol{E}}_s\boldsymbol{W}\hat{\boldsymbol{E}}_s^*\right\} \tag{10}$$

with $\boldsymbol{W} = \left(\boldsymbol{\Lambda}_s - \sigma^2\boldsymbol{I}\right)$ for deterministic ML and $\boldsymbol{W} = \boldsymbol{I}$ for MD-MUSIC. In the following discussion we consider arbitrary $\boldsymbol{W}$ which is Hermitian and positive definite.

1) Consistency: To be discussed in the lecture.

2) Asymptotic Distribution: To be discussed in the lecture.

Note: The $d^{'}$ largest eigenvectors of $\hat{\boldsymbol{R}}$ are asymptotically Gaussian distributed with means and covariances given by

$$\mathbb{E}\left[\hat{\boldsymbol{e}}_k\right] = \boldsymbol{e}_k + O\left(\frac{1}{N}\right) \tag{11}$$

$$\mathbb{E}\left[\left(\hat{\boldsymbol{e}}_k-\mathbb{E}\left[\hat{\boldsymbol{e}}_k\right]\right)\left(\hat{\boldsymbol{e}}_l-\mathbb{E}\left[\hat{\boldsymbol{e}}_l\right]\right)^*\right]=\frac{\delta_{kl}}{N}\boldsymbol{V}^{[k]}=\frac{\delta_{kl}\lambda_k}{N}\sum_{i=1,i\neq k}\frac{\lambda_i}{\left(\lambda_i-\lambda_k\right)^2}\boldsymbol{e}_i\boldsymbol{e}_i^*+O\left(\frac{1}{N}\right) \tag{12}$$

$$\mathbb{E}\left[\left(\hat{\boldsymbol{e}}_k-\mathbb{E}\left[\hat{\boldsymbol{e}}_k\right]\right)\left(\hat{\boldsymbol{e}}_l-\mathbb{E}\left[\hat{\boldsymbol{e}}_l\right]\right)^T\right]=\left(1-\delta_{kl}\right)\frac{-\lambda_k\lambda_l}{N\left(\lambda_k-\lambda_l\right)^2}\boldsymbol{e}_l\boldsymbol{e}_k^T+O\left(\frac{1}{N}\right) \tag{13}$$

C. Derivatives of a Projection Matrix

- $\boldsymbol{P}_{\boldsymbol{A}}=\boldsymbol{A}\boldsymbol{A}^\dagger=\boldsymbol{A}\left(\boldsymbol{A}^*\boldsymbol{A}\right)^{-1}\boldsymbol{A}^*=\boldsymbol{P}_{\boldsymbol{A}}^*$
- $\boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{A}=\boldsymbol{A}, \boldsymbol{P}_{\boldsymbol{A}}\boldsymbol{P}_{\boldsymbol{A}}=\boldsymbol{P}_{\boldsymbol{A}}$.
- $\boldsymbol{P}_{\boldsymbol{A}}^\perp=\boldsymbol{I}-\boldsymbol{P}_{\boldsymbol{A}}$ and $\boldsymbol{P}_{\boldsymbol{A}}^\perp\boldsymbol{A}=\boldsymbol{0}$
- In the following, we drop the subscript $\boldsymbol{A}$ for $\boldsymbol{P}_{\boldsymbol{A}}$. Note that the the derivations given below are slightly different from the ones in [1] and are taken from [2]. Starting with

$$\boldsymbol{P}_\eta=\frac{d\boldsymbol{P}}{d\theta_\eta}=\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\boldsymbol{A}\boldsymbol{A}_\eta^\dagger \tag{14}$$

and

$$\boldsymbol{A}_\eta=\frac{d\boldsymbol{A}}{d\theta_\eta}=\frac{d\boldsymbol{P}\boldsymbol{A}}{d\theta_\eta}=\boldsymbol{P}_\eta\boldsymbol{A}+\boldsymbol{P}\boldsymbol{A}_\eta\Longrightarrow\boldsymbol{P}_\eta\boldsymbol{A}=\boldsymbol{P}^\perp\boldsymbol{A}_\eta \tag{15}$$

Since $\boldsymbol{P}=\boldsymbol{A}\boldsymbol{A}^\dagger$, using (15), we have

$$\boldsymbol{P}_\eta\boldsymbol{P}=\boldsymbol{P}_\eta\boldsymbol{A}\boldsymbol{A}^\dagger=\boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger \tag{16}$$

We also have that

$$\left(\boldsymbol{P}_\eta\boldsymbol{P}\right)^*=\boldsymbol{P}\boldsymbol{P}_\eta \tag{17}$$

Now using that $\boldsymbol{P}=\boldsymbol{P}\boldsymbol{P}$, we have

$$\boldsymbol{P}_\eta=\frac{d\boldsymbol{P}}{d\theta_\eta}=\frac{d\left(\boldsymbol{P}\boldsymbol{P}\right)}{d\theta_\eta}=\boldsymbol{P}_\eta\boldsymbol{P}+\boldsymbol{P}\boldsymbol{P}_\eta=\boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\left(\boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger\right)^* \tag{18}$$

hence from (14) and (18), we have

$$\begin{aligned}\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\boldsymbol{A}\boldsymbol{A}_\eta^\dagger &= \boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\left(\boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger\right)^*\\ \Longrightarrow\boldsymbol{A}\boldsymbol{A}_\eta^\dagger &= \boldsymbol{P}^\perp\boldsymbol{A}_\eta\boldsymbol{A}^\dagger-\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\boldsymbol{A}^{\dagger*}\boldsymbol{A}_\eta^*\boldsymbol{P}^\perp\\ &= -\boldsymbol{P}\boldsymbol{A}_\eta\boldsymbol{A}^\dagger+\boldsymbol{A}^{\dagger*}\boldsymbol{A}_\eta^*\boldsymbol{P}^\perp\end{aligned}$$

$$= -\boldsymbol{P}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} + \boldsymbol{A}\left(\boldsymbol{A}^{*}\boldsymbol{A}\right)^{-1}\boldsymbol{A}_{\eta}^{*}\boldsymbol{P}^{\perp} \tag{19}$$

From (19), we can write (multiplying with $\boldsymbol{A}^{\dagger}$ from the left side)

$$\begin{aligned} \boldsymbol{A}_{\eta}^{\dagger} &= \boldsymbol{A}^{\dagger}\boldsymbol{A}\left(\boldsymbol{A}^{*}\boldsymbol{A}\right)^{-1}\boldsymbol{A}_{\eta}^{*}\boldsymbol{P}^{\perp} - \boldsymbol{A}^{\dagger}\boldsymbol{P}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} \\ &= \left(\boldsymbol{A}^{*}\boldsymbol{A}\right)^{-1}\boldsymbol{A}_{\eta}^{*}\boldsymbol{P}^{\perp} - \boldsymbol{A}^{\dagger}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} \end{aligned} \tag{20}$$

- Note that $\text{tr}\left(\boldsymbol{P}_{\eta}\right) = 0$.
- From (18), we can write

$$\boldsymbol{P}_{\eta\xi} = \boldsymbol{P}_{\xi}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta\xi}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}_{\xi}^{\dagger} + \left(\boldsymbol{P}_{\xi}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta\xi}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}_{\xi}^{\dagger}\right)^{*} \tag{21}$$

- Noting that $\boldsymbol{P}_{\xi}^{\perp} = -\boldsymbol{P}_{\xi}$ and using (18) and (20) in (21) , we get

$$\begin{aligned} \boldsymbol{P}_{\eta\xi} &= -\boldsymbol{P}^{\perp}\boldsymbol{A}_{\xi}\boldsymbol{A}^{\dagger}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} - \boldsymbol{A}^{\dagger *}\boldsymbol{A}_{\xi}^{*}\boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta\xi}\boldsymbol{A}^{\dagger} + \boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta}\left(\boldsymbol{A}^{*}\boldsymbol{A}\right)^{-1}\boldsymbol{A}_{\xi}^{*}\boldsymbol{P}^{\perp} \\ &\quad -\boldsymbol{P}^{\perp}\boldsymbol{A}_{\eta}\boldsymbol{A}^{\dagger}\boldsymbol{A}_{\xi}\boldsymbol{A}^{\dagger} + (\cdots)^{*} \end{aligned} \tag{22}$$

  where $(\cdots)^{*}$ means same expression appears with Hermitian transpose.

D. Some tricks used in the paper

1) $\text{tr}\left(\boldsymbol{ABC}\right) = \text{tr}\left(\boldsymbol{CAB}\right) = \text{tr}\left(\boldsymbol{BCA}\right)$
2) Since $\mathcal{R}\left(\boldsymbol{E}_{s}\right) \subseteq \mathcal{R}\left(\boldsymbol{A}\right)$, $\boldsymbol{P}_{\boldsymbol{A}}^{\perp}\boldsymbol{E}_{s} = \boldsymbol{0} \Longrightarrow \boldsymbol{E}_{s}^{*}\boldsymbol{P}_{\boldsymbol{A}}^{\perp} = \boldsymbol{0}$
3) If $\boldsymbol{B}$ is Hermitian, and $\boldsymbol{A}$ has the same dimensions as $\boldsymbol{B}$, we have

$$\text{tr}\left\{\left(\boldsymbol{A} + \boldsymbol{A}^{*}\right)\boldsymbol{B}\right\} = \text{tr}\left\{\boldsymbol{AB}\right\} + \text{tr}\left\{\boldsymbol{A}^{*}\boldsymbol{B}\right\} = 2\ \text{Re}\left[\text{tr}\left\{\boldsymbol{BA}\right\}\right] \tag{23}$$

4) Let $a_k = a_{kr} + ja_{ki}$ denote a complex number with a_{kr} and a_{ki} being the real and imaginary parts of a_k. Suppose $\{a_k\}_{k=1}^{d}$ and $\{b_k\}_{k=1}^{d}$ are sets of complex numbers, then we have

$$2\ \text{Re}\left\{\sum_{k} a_k\right\} .2\ \text{Re}\left\{\sum_{k} b_k\right\} = 4\left\{\sum_{k}\sum_{l} a_{kr}b_{lr}\right\} \tag{24}$$

   We can write (24) equivalently as

$$2\ \text{Re}\left\{\sum_{k} a_k\right\} .2\ \text{Re}\left\{\sum_{k} b_k\right\} = 2\ \text{Re}\left\{\sum_{k}\sum_{l}\left(a_k b_l + a_k^{*} b_l\right)\right\} = 4\left\{\sum_{k}\sum_{l} a_{kr}b_{lr}\right\} \tag{25}$$

5) When $\boldsymbol{S} = \mathbb{E}\left[\boldsymbol{s}\left(t\right)\boldsymbol{s}^{*}\left(t\right)\right] = \lim_{n\to\infty}\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{s}\left(t\right)\boldsymbol{s}^{*}\left(t\right)$ is full rank (d). Then $\boldsymbol{E}_s^{\dagger} = \boldsymbol{E}_s^{*}$ and $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{E}_s\boldsymbol{E}_s^{*}$.

Proof: Note that when $\boldsymbol{S}$ is full rank, $\mathcal{R}\left(\boldsymbol{A}\right) = \mathcal{R}\left(\boldsymbol{E}_s\right)$. Where $\boldsymbol{E}_s$ and $\boldsymbol{A}$ are $m \times d$ matrices. This implies that there exists a invertible $d \times d$ matrix $\boldsymbol{Z}$ such that

$$\boldsymbol{A} = \boldsymbol{E}_s\boldsymbol{Z} \tag{26}$$

.

Noting that

$$\boldsymbol{R} = \mathbb{E}\left[\boldsymbol{x}\left(t\right)\boldsymbol{x}^{*}\left(t\right)\right] = \lim_{n\to\infty}\frac{1}{N}\sum_{t=1}^{N}\boldsymbol{x}\left(t\right)\boldsymbol{x}^{*}\left(t\right) = \boldsymbol{A}\boldsymbol{S}\boldsymbol{A}^{*} + \sigma^{2}\boldsymbol{I}_m \tag{27}$$

has eigendecomposition as

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{E}_s & \boldsymbol{E}_n \end{bmatrix}\begin{bmatrix} \boldsymbol{\Lambda}_s & \\ & \sigma^{2}\boldsymbol{I} \end{bmatrix}\begin{bmatrix} \boldsymbol{E}_s^{*} \\ \boldsymbol{E}_n^{*} \end{bmatrix} = \boldsymbol{E}_s\boldsymbol{\Lambda}_s\boldsymbol{E}_s^{*} + \sigma^{2}\boldsymbol{E}_n\boldsymbol{E}_n^{*} \tag{28}$$

with

$$\boldsymbol{E}_s\boldsymbol{E}_s^{*} + \boldsymbol{E}_n\boldsymbol{E}_n^{*} = \boldsymbol{I}_m \tag{29}$$

and $\boldsymbol{E}_s^{*}\boldsymbol{E}_s = \boldsymbol{I}_d$ and $\boldsymbol{E}_n^{*}\boldsymbol{E}_n = \boldsymbol{I}_{m-d}$. From which we have $\boldsymbol{E}_s^{\dagger} = \left(\boldsymbol{E}_s^{*}\boldsymbol{E}_s\right)^{-1}\boldsymbol{E}_s^{*} = \boldsymbol{E}_s^{*}$.

Now $\boldsymbol{P}_{\boldsymbol{A}} = \boldsymbol{A}\left(\boldsymbol{A}^{*}\boldsymbol{A}\right)^{-1}\boldsymbol{A}^{*} = \boldsymbol{E}_s\boldsymbol{Z}\left(\boldsymbol{Z}^{*}\boldsymbol{E}_s^{*}\boldsymbol{E}_s\boldsymbol{Z}\right)^{-1}\boldsymbol{Z}^{*}\boldsymbol{E}_s^{*} = \boldsymbol{E}_s\boldsymbol{E}_s^{*}$.

III. Source Localization Application

The system model for the source localization problem from [4] is presented below.

- A source $\mathcal{S}$ is located at $(x_0, y_0) \in \mathbb{R}^2$.
- A set $\mathcal{A} \triangleq \{\mathcal{A}_1, \ldots, \mathcal{A}_M\}$ of M anchors are used for locating the source. Anchor $\mathcal{A}_m$ is located at $(x_m, y_m) \in \mathbb{R}^2$.
- The true distance between $\mathcal{S}$ and $\mathcal{A}_m$ is denoted as

$$d_m = \sqrt{\left(x_0 - x_m\right)^2 + \left(y_0 - y_m\right)^2} \tag{30}$$

- N independent measurements $r_{m,n}$ of each distance d_m are available

$$r_{m,n} = d_m + e_{m,n} = \sqrt{\left(x_0 - x_m\right)^2 + \left(y_0 - y_m\right)^2} + e_{m,n}, \quad n = 1, 2, \ldots, N \tag{31}$$

where $e_{m,n}$ are measurement errors.

- We consider two source localization methods
  - Least squares (LS)
  - Squared-range least squares (SR-LS)
- LS:

$$(\hat{x}_{LS}, \hat{y}_{LS}) = \arg\min_{x,y} \sum_{n=1}^{N} \sum_{m=1}^{N} \left(r_{m,n} - \sqrt{(x - x_m)^2 + (y - y_m)^2} \right)^2 \tag{32}$$

- We can show that the problem in (32) can be equivalently be written as

$$f_{LS}(x, y) \triangleq \sum_{m=1}^{M} \left(r_m - \sqrt{(x - x_m)^2 + (y - y_m)^2} \right)^2 \tag{33}$$

  where $r_m \triangleq \frac{1}{N} \sum_{n=1}^{N} r_{m,n} = \sqrt{(x_0 - x_m)^2 + (y_0 - y_m)^2} + e_m$ and $e_m = \frac{1}{N} \sum_{n=1}^{N} e_{m,n}$ is the averaged noise.
- LS is equivalent to ML if $e_{m,n}$ are i.i.d zero mean Gaussian distributed.
- SR-LS:

$$f_{SR-LS}(x, y) \triangleq \sum_{m=1}^{M} \left(r_m^2 - \left((x - x_m)^2 + (y - y_m)^2 \right) \right)^2 \tag{34}$$

- For $M = 2$, LS and SR-LS are equivalent in all cases of practical interest.
- For $M > 2$, LS and SR-LS are not equivalent.
- Assuming that $e_{m,n}$ are independent with zero mean, $\mathbb{E}\left[e_{m,n}^2\right] = \sigma^2$, $\mathbb{E}\left[e_{m,n}^3\right] = 0$, and $\mathbb{E}\left[e_{m,n}^4\right] = \alpha\sigma^4$ for some constant α. For Gaussian case $\alpha = 3$.
- This results in $\mathbb{E}(e_m) = 0, \mathbb{E}(e_m^2) = \frac{\sigma^2}{N}$ and $\mathbb{E}(e_m^4) = \frac{\sigma^4}{N^2}\left(3 + \frac{\alpha - 3}{N}\right)$.
- For convenience, without loss of generality, we assume that true location of source $\mathcal{S}$ is $(x_0, y_0) = (0, 0)$.
- Some notations

$$X_a(\mathcal{A}) \triangleq \sum_{m=1}^{M} \frac{x_m^2}{x_m^2 + y_m^2}, Y_a(\mathcal{A}) \triangleq \sum_{m=1}^{M} \frac{y_m^2}{x_m^2 + y_m^2}, Z_a(\mathcal{A}) \triangleq \sum_{m=1}^{M} \frac{x_m y_m}{x_m^2 + y_m^2}$$

$$X_b(\mathcal{A}) \triangleq \sum_{m=1}^{M} x_m^2, Y_b(\mathcal{A}) \triangleq \sum_{m=1}^{M} y_m^2, Z_b(\mathcal{A}) \triangleq \sum_{m=1}^{M} x_m y_m$$

$$X_c\left(\mathcal{A}\right) \triangleq \sum_{m=1}^{M} x_m^2\left(x_m^2+y_m^2\right), Y_c\left(\mathcal{A}\right) \triangleq \sum_{m=1}^{M} y_m^2\left(x_m^2+y_m^2\right), Z_c\left(\mathcal{A}\right) \triangleq \sum_{m=1}^{M} x_m y_m\left(x_m^2+y_m^2\right)$$

REFERENCES

[1] M. Viberg and B. Ottersten, "Sensor array processing based on subspace fitting," *IEEE Trans. Sig. Proc.*, vol. 39, pp. 1110-1121, May 1991.

[2] G. Golub and V. Pereyra, "The differentiation of pseudoinverses and nonlinear least squares problems whose variables separate," *SIAM J. Numer. Anal.*, vol. 10, pp. 413-432, 1973.

[3] J. A. Cadzow, "A high resolution direction-of-arrival algorithm for narrow-band coherent and incoherent sources," *IEEE Trans. Acoust., Speech, Signal Processing*, vol. 36, pp. 965-979, July 1988.

[4] E. G. Larsson and D. Danev, "Accuracy comparison of LS amd squared-range LS for source localization," IEEE Trans. Sig. Proc., vol. 58, pp. 916-923, Feb 2010.